IOP Publishing Journal of Neural Engineering
J. Neural Eng. **11** (2014) 000000 (14pp) UNCORRECTED PROOF

Closed-loop optical neural stimulation based on a 32-channel low-noise recording system with online spike sorting

T K T Nguyen[1,2], Z Navratilova[1,3], H Cabral[1,3], L Wang[1,4], G Gielen[1,2], F P Battaglia[3,5] and C Bartic[1,4]

Q1 [1] Imec, Belgium
[2] Department of Electrical Engineering, Katholieke Universiteit Leuven, Belgium
[3] Department of Neuroinformatics and Neurophysiology Donders Centre for Neuroscience, Radboud Universiteit Nijmengen, Netherlands
[4] Department of Physics and Astronomy, Katholieke Universiteit Leuven, Belgium
[5] VIB, Belgium

E-mail: carmen.bartic@fys.kuleuven.be

Received 4 October 2013, revised 27 March 2014
Accepted for publication 9 April 2014
Published DD MM 2014

Abstract

Objective. Closed-loop operation of neuro-electronic systems is desirable for both scientific and clinical (neuroprosthesis) applications. Integrating optical stimulation with recording capability further enhances the selectivity of neural stimulation. We have developed a system enabling the local delivery of optical stimuli and the simultaneous electrical measuring of the neural activities in a closed-loop approach. *Approach.* The signal analysis is performed online through the implementation of a template matching algorithm. The system performance is demonstrated with the recorded data and in awake rats. *Main results.* Specifically, the neural activities are simultaneously recorded, detected, classified online (through spike sorting) from 32 channels, and used to trigger a light emitting diodes light source using generated TTL signals. *Significance.* A total processing time of 8 ms is achieved, suitable for optogenetic studies of brain mechanisms online.

Keywords: closed-loop operation, real-time monitoring, online signal processing, optogenetics, neural stimulation

Q2 (Some figures may appear in colour only in the online journal)

1. Introduction

Electrical stimulation has been used in many therapeutic applications for the restoration of lost neurological functions (Peckham and Knutson 2005). Yet the stimulation is non-selective, and a large population of cells is activated or inhibited due to the complex current pathways in the tissue. Simultaneous electrical stimulation and recording would provide more information on brain dynamics but suffer from the electrical stimulation artifacts (Hashimoto *et al* 2003, Jensen and Durand 2009, Sunderam *et al* 2010). On the other hand, the optical stimulation of brain tissue, where light-sensitive proteins (e.g. chan-nelrhodopsin-2 (ChR2)) are selectively expressed in specific cell types using genetic technologies, has recently been reported as a method to achieve the stimulation of neural activities with sub-millisecond temporal resolution and cell-type specificity (Zhang *et al* 2009, Zhang and Oertner 2007, Gradinaru *et al* 2009, Scanziani and Husser 2009). With this technique, which has been validated both *in vitro* and *in vivo* (Huber *et al* 2008, Zhang and Oertner 2007), light is delivered using various schemes, including light emitting diodes (LED), laser-coupled optics or direct epicranial LED implantation (Royer *et al* 2010, Stark *et al* 2012).

The combination of neural stimulation (be it electrical or optical) and electrical recordings from multi-electrode

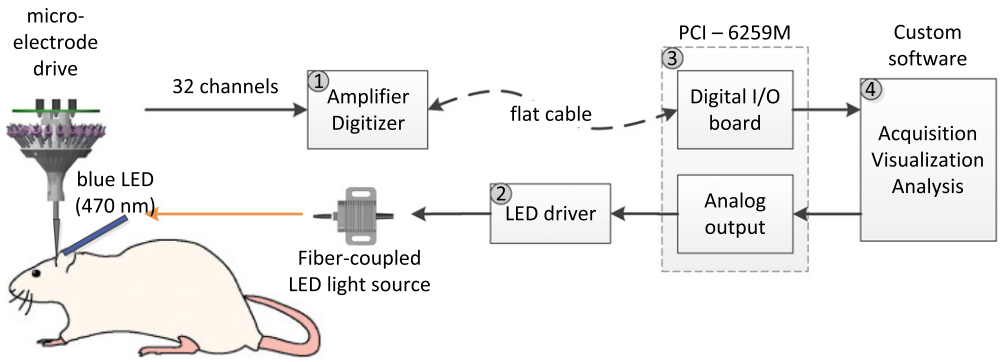

Figure 1. The implemented closed-loop system architecture. (1) A headstage acquires neural signals from 32 electrodes, amplifies, multiplexes, digitizes and sends them to the digital I/O board. (2) A LED driver controls the fiber-coupled LED light source. (3) A data acquisition card (National Instruments PCI-6259M) interfaces the digital and analogue inputs/outputs. (4) A custom software controls the acquisition, visualization and analysis of the recorded signals.

arrays has been performed in two manners: (1) open-loop and (2) closed-loop operation. In a typical open-loop experiment, the stimulation parameters are set beforehand, and the responses are measured correspondingly. The deep-brain stimulation currently used to treat the Parkinson disease is a typical example of this operation. In contrast, in a closed-loop operation, the responses of the nervous system itself determine the next stimulus. Studies performed in the closed-loop operation promise to provide more information about the neural dynamics and lead to the more effective treatment of nervous system disorders (Fountas and Smith 2007, Cunningham *et al* 2010, Liang *et al* 2011, Rosin *et al* 2011, Wagenaar *et al* 2005, Cerf *et al* 2010, Potter *et al* 2006, Santos *et al* 2011). Typical examples include studying the mechanisms of a pacemaker's synchronized firing in a demand-controlled way (Tass 2002, 2003, Tass *et al* 2003, Hauptmann *et al* 2007, Popovych *et al* 2005); altering functional connectivity and motor behavior through single-electrode stimulation based on extracellular recordings in the monkey cortex (Jackson and Fetz1 2006); regulating the firing activity of the medial temporal lobe neurons through decoding the visibility ratio every 100 ms (Cerf *et al* 2010); etc.

The closed-loop operation is advantageous when neural selectivity is desired. The stimulation signal can be triggered online, whenever a particular neural activity pattern is detected (Sahin *et al* 2000, Venkatraman *et al* 2009). This requires online spike sorting. Yet this often needs to be sacrificed for the speed of real-time decoding or visualization (∼100 ms) (Cerf *et al* 2010, Zrenner *et al* 2010). Few algorithms have been realized and implemented in a real-time closed-loop stimulation and recording system (Guger *et al* 2011), despite the existence of several algorithms for online processing capability (Oliynyk *et al* 2012, Franke *et al* 2010, Rutishauser *et al* 2006, Chandra and Optican 1997). Additionally, the controlling frequency for both the recording system and the device output is low in those cases, e.g. 48 Hz (∼20 ms) reconstruction time (Guger *et al* 2011).

In this paper we describe a system allowing electrical recording and activity-based optical stimulation. This is interfaced with a microdrive that can host up to 24 tetrodes and 2 optical fibers. The prototype is closed-loop operated by implementing the template matching as a spike sorting method. Our implementation is fast enough to process at least 8 integrated tetrodes (32 channels) with a controlling time of 8 ms. In particular, the signal processing sequences including acquisition, detection, classification, and pulse delivery are operated online. The processing accuracy and speed of the integrated software are demonstrated first with recorded data and later with recordings in awake animals. The LED light is delivered through the fiber to trigger neural cells based on the sorting results. The resulting prototype offers a platform to enhance the study of brain dynamics in optogenetic experiments.

2. System architecture

The main objectives for this first-generation closed-loop system were (1) to produce a low-noise, scalable system with at least 32 channels, and (2) to integrate online neural signal analysis.

The overall system is shown in figure 1, and it consists of four main components: (1) a headstage that amplifies, multiplexes and digitally transmits the signals to the digital I/O board, (2) a LED driver that controls the brightness of the LED through the current pulses, (3) a data acquisition card (PCI-6259M) that provides digital I/O and analogue output pulses and (4) a custom software controlling the acquisition, visualization and analysis of the recorded signals. The headstage and the LED light sources (optic fibers) connect to the implanted microwires via a Neuralynx electrode interface and a micro-drive. The board provides the signal connection between electrode wires and the drive. The drive can host up to 24 tetrodes, each consisting of a twisted bundle of four or eight microwires and each coupled to a guiding unit with a micro screw, allowing vertical displacement of the tetrode. The fabrication and assembly of such a microdrive array is described in Kloosterman *et al* (2009).

2.1. Amplifier board

The amplifier board is designed for the Intan Technologies RHA2132 amplifier chip (Intan technologies html) (see figure 2).

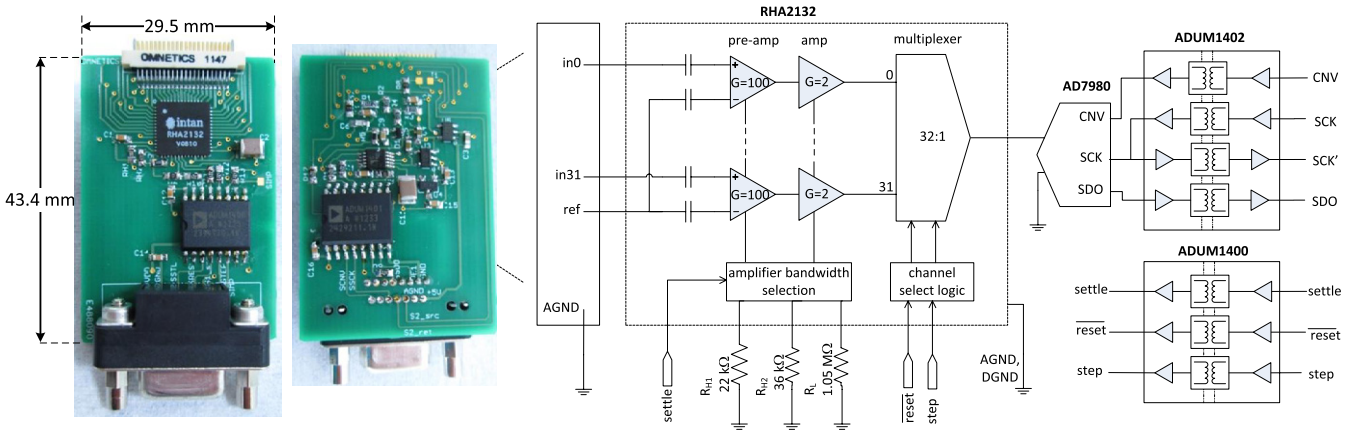

Figure 2. A 32-channel recording system. (Left) the assembled PCB, built on 29.5 mm × 43.4 mm, contains a miniature connector, a packaged amplifier, an ADC and digital isolators. (Right) schematic illustrating the signal path. Pre-amplification = 100×, amplification = 2×, filter bandwidth = 0.2–5000 Hz, multiplexing ratio = 32:1, 16-bit ADC and digital isolators.

The 32-channel amplifier board interfaces with the micro-electrode drive through a 44-pin connector (nano-series, Omnetics).

Neural signals are recorded against a common reference electrode. The ground is connected to the drive ground. Although the amplifier gain is fixed at 200×, the signal bandwidth can be set by external resistors, resulting in a bandwidth ranging from 0.2 Hz to 5 kHz. The amplifier is DC-coupled, with a gain of 1 at 0 Hz. The multiplexed signals are digitized with 16-bit resolution (AD7980) before sending them to the digital I/O board.

The integrated multiplexers permit sampling speeds of up to 31.25 kSps (kilosamples per second) per channel. The maximum aggregate sampling rate of 10 MHz of the DAQ device (PCI-6259M) however results in a TTL minimum pulse width of 100 ns (=1/10 MHz). Considering the requirements of the ADC conversion and acquisition time, we chose to set a fixed cycle time of 2500 ns (i.e. 900 ns of conversion and 16 × 100 ns of acquisition of 16-bit data samples), limiting the sampling speed to 400 kSps (=1/2500 ns) or 12.5 kSps per channel. The high digital transmission speed of 10 MHz can cause several side effects, such as ringing, crosstalk, reflections and ground bounce. To ensure the signal integrity, impedance bridging source termination and line drivers are implemented. In the next generation, a micro-controller or field-programmable gate array (FPGA) can be placed on the amplifier board to further reduce the TTL minimum pulse width, synchronize the acquisition and remove the transmission of conversion TTL pulses, which ultimately increases the sampling rate.

The amplifier board is connected to the acquisition card by a flat ribbon, 14-conductor, unshielded cable. Although the cable is lighter compared to the corresponding shielded one, its weigh and flexibility would still be a bottleneck regarding the constrained movement of the animals. Yet this commercial cable with standard impedance matching facilitated the implementation of our first closed-loop prototype. The 14 conductors alternate between ground and signal, with the ground minimizing the crosstalk between the channels. The seven signal-carrying lines are: (1) digital input channel selector, commanding the amplifier to sample the next channel when this line is activated; (2) digital channel reset, which commands the amplifier to return to the first input channel; (3) digital settle, for rapidly discharging all capacitors in the front-end amplifiers to ground in the event of amplifier saturation; (4) ADC conversion signal, which initiates a conversion and outputs the most significant bits on its falling edge when the conversion is complete; (5) ADC clock signal, which outputs the remaining data bits at subsequent falling edges in the acquisition phase; (6) the serial digital output of the ADC, returning the converted acquired signal; and (7) the return ADC clock signal, which ensures accurate signal acquisition at the computer's side. For supporting a larger number of channels (scalability) while maintaining low wiring requirements (board compactness), we would only need to implement ADC outputs. The controlling signals can be shared among the different amplifier boards.

To further ensure a low-noise architecture, we have placed digital isolators (ADUM1400 series) to decouple the ground of the amplifier board from the noisy computer ground. The amplifier board operates on two battery-supplied power sources: a regulated +3V and +5V. A 2.5V reference voltage (ADP1710) is used for the ADC conversion.

2.2. LED driver

A simple high-power LED driver supporting pulse-width modulation has been built to control the brightness of the LEDs. The analogue output of the acquisition card delivers a pulsed signal with pre-defined stimulation duration to the transistor. The resulting output current then modulates the brightness of the LED.

2.3. Controlling software

A custom software has been implemented to control the acquisition, visualization and signal analysis. It consists of two parts: (1) the hardware-timed acquisition, and (2) the online signal processing.

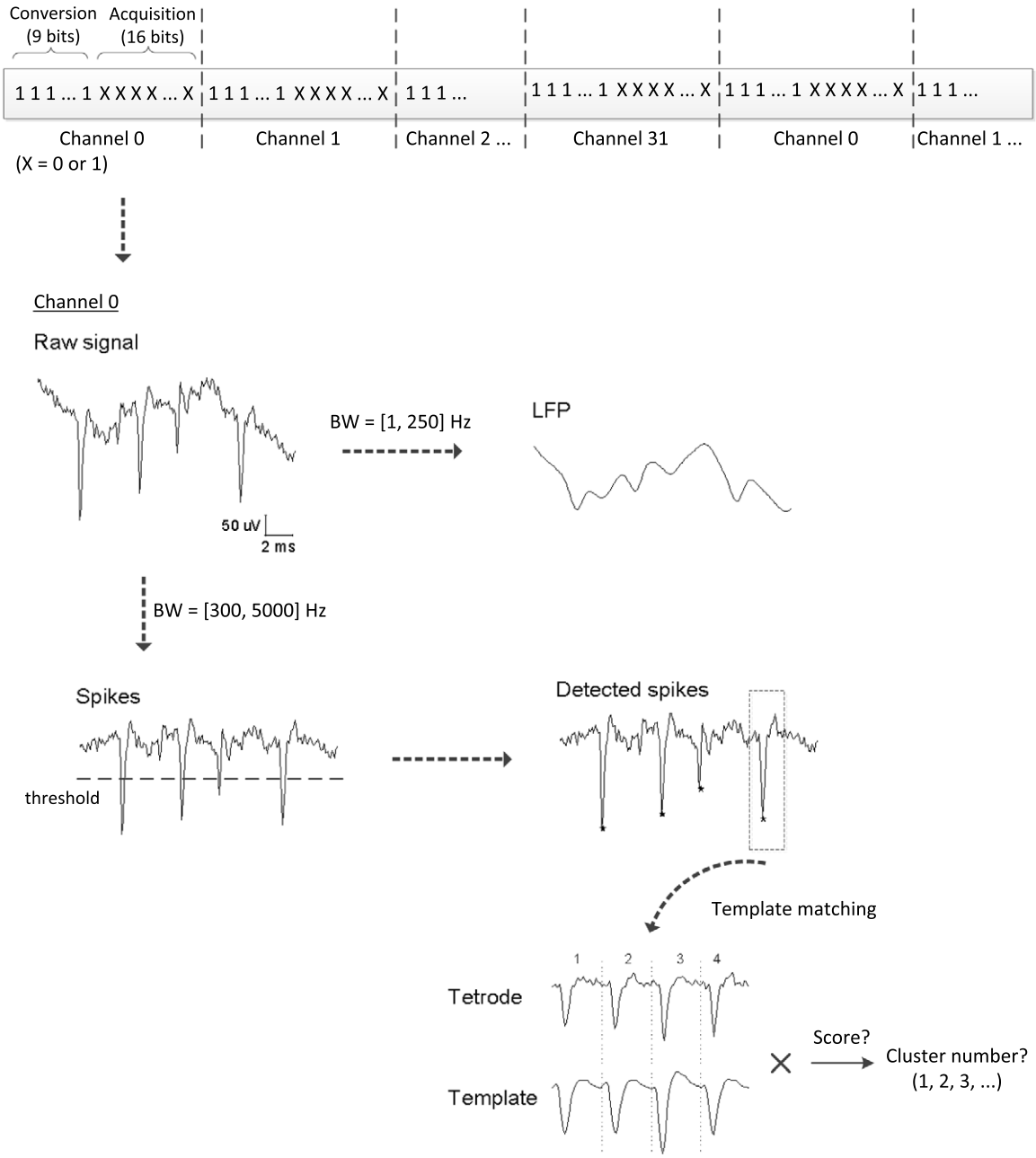

Figure 3. Signal processing steps: a raw multiplexed digital signal sequence is acquired, demultiplexed, then converted to an analogue signal by multiplying with a 16-bit binary code. It is then filtered by an elliptic filter of 300 Hz to 5 kHz, and detected by amplitude thresholding (the calculation for the noise level is described in Quiroga *et al* (2004)). The detected spikes are then compared with existing templates and classified based on the highest score.

Hardware-timed acquisition: the board is operated with software created in LabVIEW (National Instruments). Data from the PCB are acquired by applying the producer–consumer design pattern. Parallel loops are broken down into two categories: high-speed data acquiring (producer loop), and data storing and analysis (consumer loop). Data queues are used to communicate data between the loops, minimizing data loss. As a result, the consumer loop (which is normally slower) processes the data at its own pace while allowing the producer loop to keep up with the hardware-timed acquisition. If the consumer loop is really slower and the buffer is full, an error message appears, breaking the acquisition. Thus, a relative-balanced producer–consumer processing time is required.

Online signal processing: the data processing sequence is illustrated in figure 3, and consists of four steps: (1) de-multiplexing and converting the binary code to an analogue voltage; (2) filtering the recorded data; (3) detecting spikes in every tetrode (4 channels); and (4) classifying the spikes.

- Signal de-multiplexing and conversion: the raw multiplexed digital data from the ADC output are de-multiplexed to separate them into their constituent

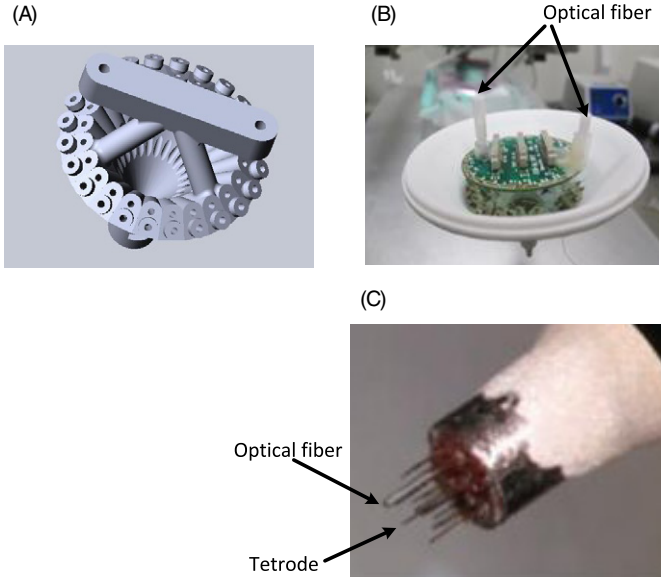

Figure 4. (A) CAD drawing of the 'split bundle' micro-drive for joint prefrontal cortex-hippocampal recordings and optical stimulation; (B) Photograph of one of the practical implementations containing the bundle microdrive loaded with tetrodes and 2 optical fibers and the PCB connecting the tetrodes to the data acquisition; (C) Cross-sectional view through the tetrode/fiber bundle of (B).

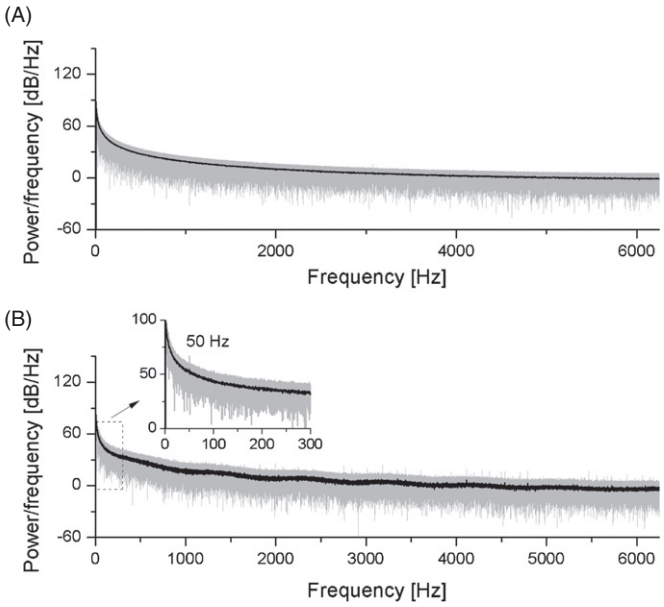

Figure 5. Power spectral density of channel 1. The black lines denote smoothed signals; shadings represent the actual density calculated by FFT. (A) Noise spectra in the system with grounded inputs, (B) Noise spectra with the electrodes implanted in the brain. The inset corresponds to the total noise spectrum (electronic, electrode and biological noise sources) in the local field potential band (∼1–300 Hz).

input channels. Each binary data stream is subsequently removed 9 overhead zeros and converted to a 16-bit data. The input-referred voltage can be derived from this value by scaling with $\times\frac{2.5}{2^{16}} \times \frac{1}{200}$ where 2.5 V is the ADC reference voltage, and 200× corresponds to the nominal gain of the amplifier.

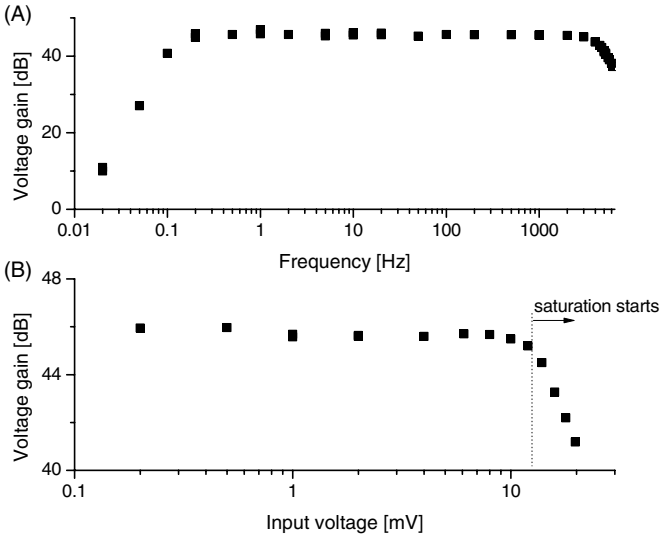

Figure 6. (A) Measured gain of a channel from 0.02 Hz to 6.2 kHz. The amplifier was configured for 0.2 Hz to 5 kHz. (B) Measured gain of a channel with a 1 kHz sine wave from 0.2 to 20 mV_{pp}.

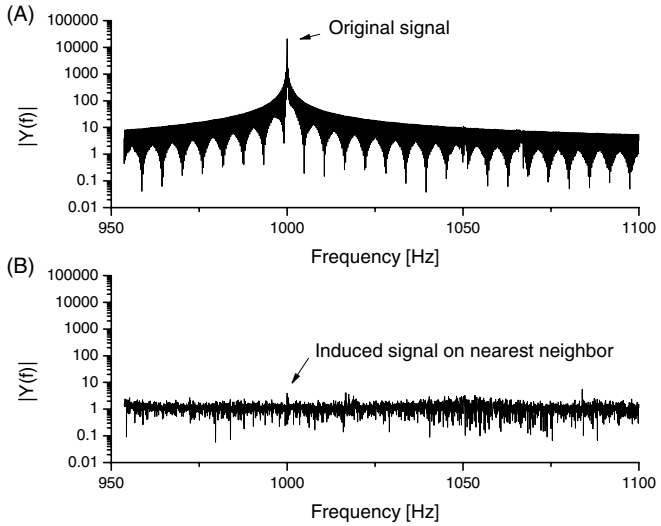

Figure 7. Single-sided amplitude spectrum of (A) original signal (10 mV_{pp}, 1 kHz sine wave), and (B) signal induced on the nearest neighboring channel. This 'worst-case' situation results in a crosstalk of −74.4 dB.

- Signal filtering: signals are then bandpass filtered with a 2 nd-order elliptic filter. While the bandwidth of the raw data is from 0.2 Hz to 5 kHz, the filtered signal for spike detection usually has a bandwidth of 300 Hz–5 kHz.
- Spike detection: spikes are detected based on the negative amplitude threshold method. The threshold (thr) is automatically calculated using the formula proposed by Quiroga *et al* (2004),

$$\text{thr} = k\sigma_n; \quad \sigma_n = \text{median}\left\{\frac{|x|}{0.6745}\right\} \qquad (1)$$

where x is the bandpass-filtered signal, k = 3, 4, 5, ... (set by the experimenter), and σ_n is an estimate of the standard deviation of the background noise. This estimate

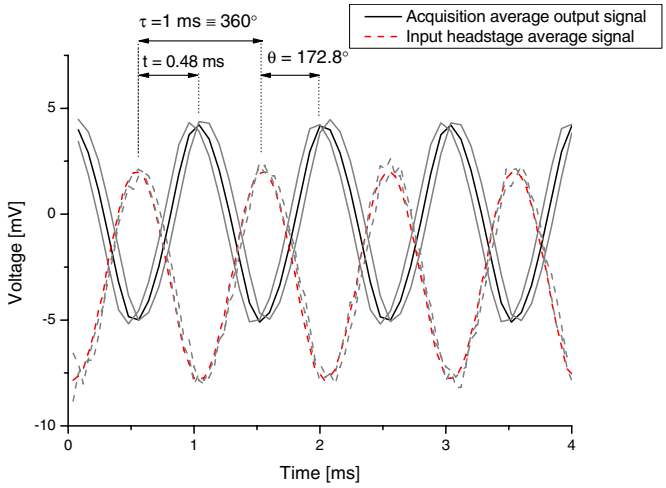

Figure 8. Phase offset between a 1 kHz reference signal (from a function generator (red, dash line)) and the signal measured with the acquisition system (black, solid line). The gray lines denote the boundaries of the corresponding signals. The negative phase denotes that the output is lagging the input.

was proven to be more effective in setting threshold values in the presence of high interferences (Quiroga *et al* 2004).

A 1 ms signal excerpt from the 4 channels of a tetrode is considered to be a set of spikes if one of its four peaks is greater than the corresponding threshold. If two spikes occur so close to each other that the time difference is smaller than the reference time (a predefined minimum time between two spikes), then the first spike is kept while the second is discarded.

For each detected spike, 32 samples (i.e. 2.6 ms) are saved for further analysis. In order to avoid misalignments, all spike sequences are taken 8 samples prior to and 24 samples post the maximum peak.

- Spike classification: spike sorting is carried out using a multi-channel template matching process. Templates are extracted from a baseline recording period at the beginning of the session. Our current template definition is based on a standard mixture-of-Gaussians expectation maximization algorithm (Harris *et al* 2000). The number of Gaussian components (i.e. putative recorded cells) is at this point left to the experimenter. Online spike sorting is then performed using the simplest form of template matching, i.e. using the dot product where single spike waveforms are compared to each template. The normalized dot product of the waveform with all templates is taken, and the spike is then assigned to the template resulting in the maximum score of correlation.

The whole signal analysis process is performed by deployed Matlab code in LabVIEW. A dynamic-link library (DLL) is first compiled from the Matlab code using the deployment tool. Then, a wrapper is written in C language to perform the Matlab compiler runtime (MCR) initialization part. Once a C DLL is generated, it can be used in LabVIEW (Wang *et al* 2012).

Table 1. Measured specifications and performances of the application-specific circuit used in these experiments.

Parameter	Value
Number of input channels	32
Number of wires needed for recording	7 + 4 (power lines)
Board size	29.5 mm × 43.3 mm
Power supply voltage	3 V & 5 V
Low-frequency cutoff	0.17 Hz
High-frequency cutoff	4.5 kHz
Input-referred noise	2.5 ± 0.24 μV_{rms}
Gain	190 V/V
Average adjacent input channel crosstalk (1 kHz)	−74 dB

Note: the reduction in the number of wires needed for recording relative to the number of inputs achieved through on-chip multiplexers.

3. Validation methods

3.1. Amplifier board characterization

Noise measurements: noise is measured under two conditions: (1) input grounded, and (2) input connected to tetrodes implanted in the brain. The measurements are carried out on 32 channels with a 12.5 kHz sampling rate per channel. The root-mean squared (rms) noise values are obtained by taking the standard deviation of each channel's voltage-time trace. No filter is applied.

Amplifier gain measurements: the nominal gain and phase responses of the system are measured by applying 1 mV_{pp} sine waves in the range from 0.02 Hz to 6.2 kHz from a signal generator (TG2000). The ADC output and the sinusoidal input signal are connected to the digital input and the analogue input of the acquisition card respectively. The voltage gain is calculated from the fast Fourier transform (FFT) of the two signals.

We have also verified that the system's gain is constant across a range of input amplitudes. We have tested this with a 1 kHz sine wave, with varying amplitudes from 0.2 to 20 mV_{pp}.

Cross-talk and phase offset measurements: cross-talk is measured by sending a 1 kHz sine wave of 1 mV_{pp} to one channel, and recording from the grounded neighboring channels.

The phase offset is measured by applying a 1 kHz sine wave to a channel, and connecting directly to the analogue input of the DAQ card. This also gives the signal propagation delay. The acquisition is controlled by the same clock (ADC acquisition clock).

Settling time: the ability to recover from large transients (artifacts) of the system is integrated into the function of the *settle* pin, which is held low in normal operation, and is pulled high momentarily to settle the amplifiers.

The actual required time to settle for all the amplifiers is measured with interferences of 50 Hz on the open inputs (i.e. not in connection with the optogenetics stimulation). An active-high digital input is delivered to the *settle* pin until

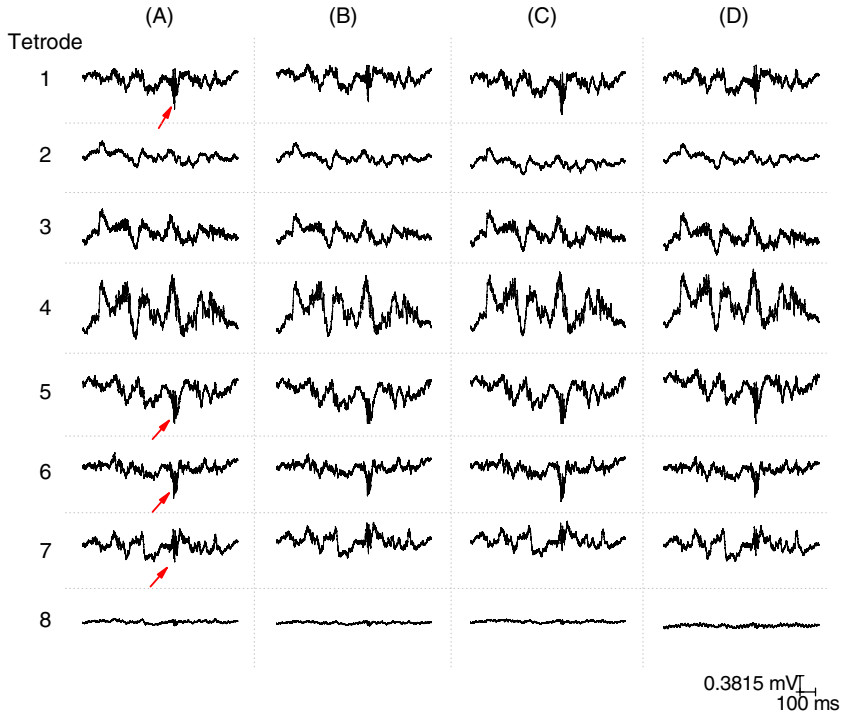

Figure 9. Simultaneous recordings from six tetrodes (24 channels) in an awake rat. The majority of sites report action potentials. Tetrode 2 and 8 are non-functional since they are open contacts. Each of the four wires (A)–(D) of a tetrode indicates relatively similar patterns. The arrows indicate putative (or probable) hippocampal sharp waves. The signal bandwidth is between 0.2 and 5000 Hz.

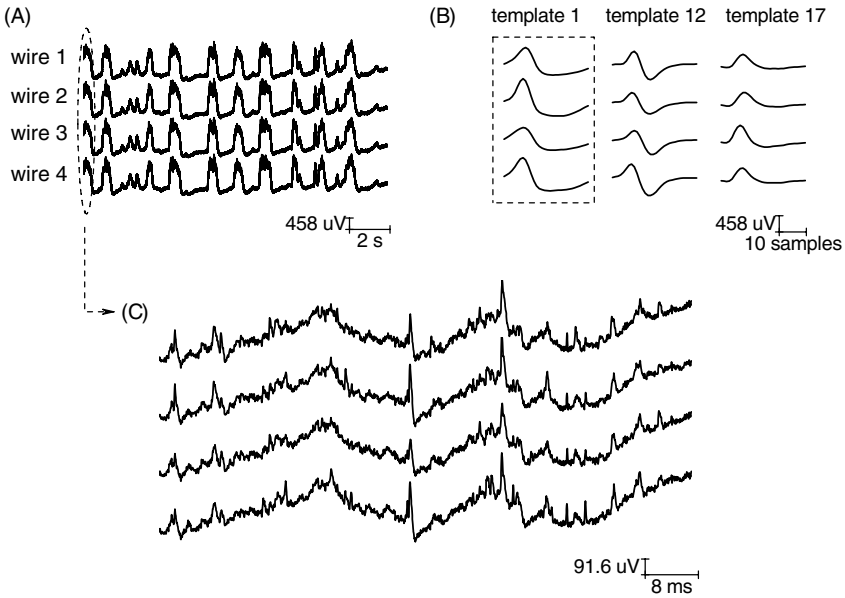

Figure 10. Recorded data set used for spike sorting. (A) A 15 s sequence of the raw data, showing cortical UP/DOWN states. (B) Examples of three template spike shapes. (C) A smaller data fragment of (A).

the large 50 Hz interferences (often saturated) disappear, and the signals are back to the baseline. During that period, the clock sequentially steps through all the channels like in normal recording.

3.2. Animal surgeries and recordings

To prove the system's functionality in awake animals, microwire array recordings have been taken from the dorsal hippocampus of adult male Sprague–Dawley rats weighing >350 gr (Charles River Laboratories; Wilmington, MA, USA). The experiments have been conducted in accordance with protocols approved by the KU Leuven animal ethics committee and in accordance with the European Communities Council Directive of November 24, 1986 (86/609/EEC). Genetic material for ChR2-mCherry was delivered to the mouse neurons using AAV under the control of the CaMKIIα promoters for wild-type mice. Rats are

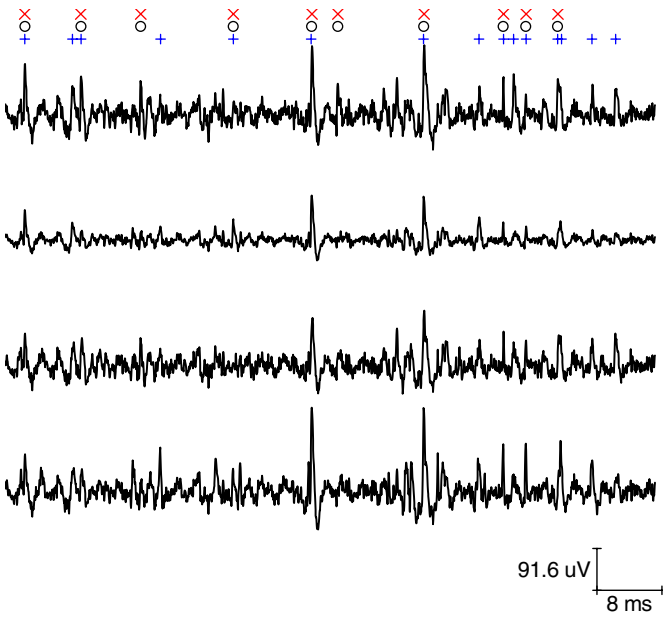

Figure 11. Bandpass-filtered signal with a 2 nd-order elliptic filter from 0.6–6 kHz. The '**x**' indicate the locations of Neuralynx detected spikes. The '**o**' are the locations of the sorted spikes. The '**+**' symbols indicate the spikes detected with a threshold of $4.5\sigma_n$, or $50.24 \pm 1.17\ \mu V_{rms}$ in this case.

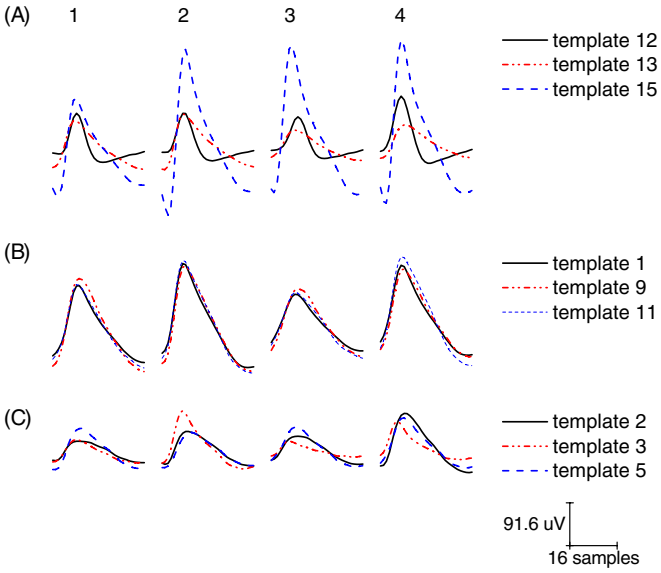

Figure 12. Examples of templates used for spike sorting. (A) Templates with visibly different features. (B) Templates with visibly similar features. (C) Other templates with several feature differences.

anesthetized with 1.5–3.0% inhaled isoflurane and are given a subcutaneous injection of buprenorphine (0.05 mg kg^{-1}) to minimize pain. A craniotomy is made over the right dorsal hippocampus, centered at 3.5 mm posterior and 2.8 mm lateral to bregma. The dura is incised with a sterile syringe needle and a microwire array (Kloosterman *et al* 2009) with 24 tetrodes is implanted.

Tetrode construction is as previously described (Wilson and McNaughton 1993): each tetrode consists of a twisted bundle of four polyimide-insulated microwires, fused and cut to create a blunt tip. The wire used for the tetrodes is a 13-μm diameter nichrome resistance wire (California Fine Wire, Grover Beach, CA) with all recording sites plated with gold simultaneously using an electroless dip-plating process (Immersion Gold CF, Transene, Danvers, MA). An example of the assembled micro-drive for electrical recording and optical stimulation used in the experiments is shown in figure 4.

The craniotomy is sealed with a bio-compatible, silicon-based elastomer (Kwik-Sil from World Precision Instruments), skull screws (Amazon Supply, US) are implanted and UV-curable cement (SDI, AUS) is applied to secure the array. The rats are returned to their normal housing, rest and recover post-operatively for 5–8 days before the recordings begin. In the meantime, tetrodes are lowered while monitoring the activity in order to attain the correct positioning.

3.3. Signal processing

Tests of online signal processing have been conducted with two sets of data: (1) neural data recorded with the Neuralynx system, and (2) data acquired and analyzed using our custom-built system.

Neural signals in the hippocampus are acquired by the implanted tetrodes using the Neuralynx system (sampling frequency of 32556 Hz/channel, bandwidth of 0.1 Hz to 9 kHz, scaling ratio of 9.1e-8). The data are pre-processed (denoising and removing interferences), amplitude threshold detected and classified by fitting a mixture of Gaussians (KlustaKwik) (Harris *et al* 2000, Kadir *et al* 2013). The spike clusters might be different when using other spike sorting programs (e.g. wavelets and super-paramagnetic clustering of the Waveclus program (Quiroga *et al* 2004)). However, this work does not focus on the robustness of the sorting process algorithms (i.e. offline analysis), but rather on the online performance of the system. Therefore, the clusters, which resulted from the KlustaKwik program, were used here. The signal bandwidth is restricted from 0.6 to 6 kHz. The extracted templates are the average spike waveforms in each cluster. A template consists of 4×32 points, where 4 is the number of channels in one tetrode; 32 points = 8 pre- + 24 post-waveform segments around the detected peak.

For the recorded data, the speed of data processing (consisting of filtering, signal detection and signal matching) has been measured under different acquisition time windows. Then, the similarities between the template-classified spikes and the fine-tuning manually classified spikes have been measured to provide a quantitative value for the algorithm's reliability.

For the data acquired with our set-up, neural signals are recorded from eight tetrodes every 8 ms. The signals are amplified, bandpass filtered over 0.2–5000 Hz, multiplexed, digitized and sampled with a clock of 10 MHz. Each acquisition consists of three activities: (1) store for later offline analysis; (2) detect and classify spikes (300–5000 Hz); and (3) deliver a TTL pulse based on the pre-defined cluster.

Table 2. Number of missed spikes and false positives between our amplitude threshold-based detection method and the Neuralynx analyzed data set

Threshold level	Detected hit rate (%)	Sorted hit rate (%)	False positives (%)
$5.5\sigma_n$	56.3 (175/295)	54.2 (150/277)	37.5 (105/280)
$5\sigma_n$	71.8 (210/295)	65.3 (181/277)	43.4 (161/371)
$4.5\sigma_n$	77.3 (228/295)	70.8 (196/277)	51.7 (244/472)

The analysis is performed with a 2 nd-order elliptic filter every 15.7-ms sequence of data for 1000 cycles.

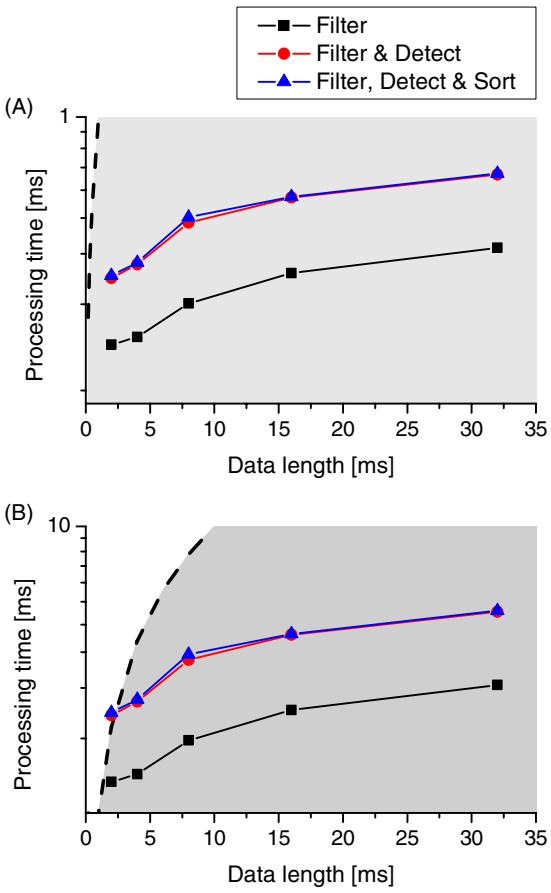

Figure 13. Processing time versus data length for (A) 4 channels, and (B) 32 channels. The gray area indicates the room and boundary for online processing time. The time in (B) is simply a scaling factor from (A), indicating the time challenge in online data processing for a larger number of recordings. The filtering time is the most time consuming when the data length is ∼2 ms. The values are the median of 1000 cycles, for a 2 nd-order elliptic filter.

4. Experimental results

4.1. Performance of the 32-channel low-noise amplifier board

The assembled PCB, measuring 29.5 mm × 43.3 mm, contains surface-mounted passive components for performing voltage stabilization, a voltage reference and regulators for controlling the amplifier and the ADC. The placement of the digital isolator and the division of power planes are performed to minimize the noise from the external interferences. Signals are amplified, filtered, multiplexed, digitized and transmitted through line drivers via a flat cable to the digital I/O board. A minimum of seven wires (channel sweep, amplifier reset, amplifier settle, ADC clock, ADC conversion, signal synchronization and ADC data output) excluding power supply is required to perform recording functions with the amplifier board. Control signals (six lines) however can be shared between multiple 32-channel headstages. Therefore, the number of lines increases only by one (the ADC output) when adding a 32-channel headstage.

Table 3. Correct hit rates of spike sorting using template matching. The results are obtained with 2 nd-order elliptic filtering, and a threshold level of $4.5\sigma_n$.

Cluster number	Total sorted spikes	Correctly classified spikes	Matching rate (%)
12	70	62	88.57
13	46	39	84.78
15	61	55	90.16
1	135	49	36.30
9	20	10	50
11	2	20	10
2	54	35	64.81
3	73	50	68.49
5	17	11	67.71

The overall measured specifications of the designed prototype are presented in table 1.

Low-noise characteristic of the recording system: Figure 5(A) shows the noise spectrum of channel 1 with grounded input. An average noise level of 2.5 ± 0.24 μV_{rms} has been measured.

With implanted tetrodes, the system also shows a low noise level (see figure 5(B)). Although 50 Hz interference is still observed, its level is negligible to the signals (see the inset in figure 5(B)). The noise level of our amplifier board is comparable with commercial systems (such as the Neuralynx).

Amplifier gain: Filter-related gain changes have been observed at frequencies near the band-pass filter cut-off frequencies. Although the signal bandwidth is set from 0.2 Hz to 5 kHz by three external resistors, we have observed a passband gain of about 45.57 dB (190 V/V), and the measured −3 dB frequencies were located at 0.17 and 4.7 kHz (see figure 6(A)).

The amplifier gain is approximately constant from 200 μV_{pp} up to 12 mV_{pp}, and is 45.64 ± 0.2 dB (figure 6(B)). This also implies that the recorded signal amplitudes are reflected linearly over this range of physiological inputs. Yet, we observed saturation occurring in the negative pulse when

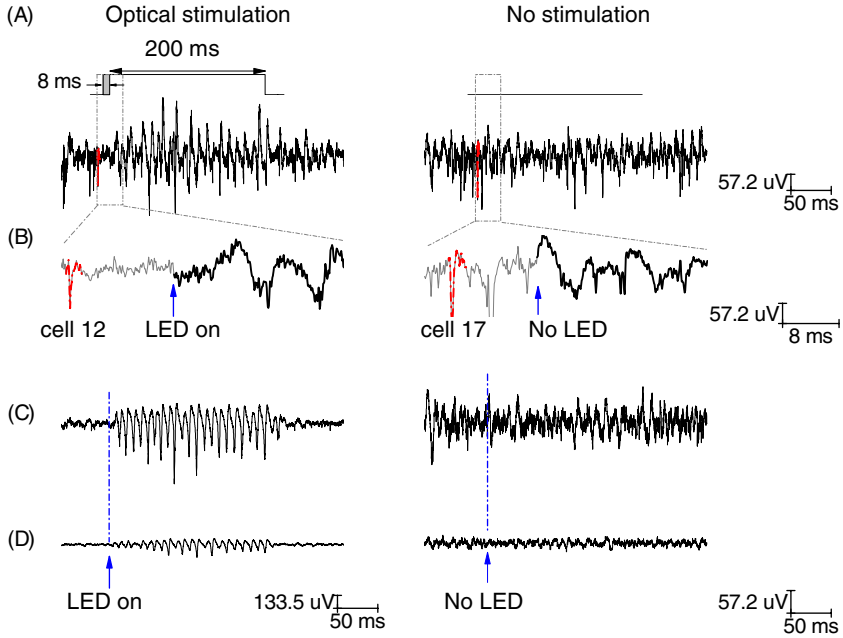

Figure 14. Examples of optical stimulation based on spike sorting. (A) A TTL pulse of 200 ms is delivered to the optical fiber whenever the spikes are classified as cluster 12 (right). Otherwise, there is no stimulation (left). The gray rectangle is the 8-ms processing after receiving the 8-ms acquisition. (B) The zoom-in of 40 ms of data (5 frame of acquisition). The sorted spikes are indicated by the dashed line. The arrow indicates the start of stimulation/no-stimulation. (C) Strong stimulation effect on nearby tetrodes. (D) Weak stimulation effect on far-away tetrodes. The displayed signals are band-pass filtered from 300 to 5000 Hz.

the input voltage is greater than 12 mV_{pp}, which indicates the presence of a negative offset in the range of -109 μV, referred to the amplifier input, which is in agreement with the RHA2132 datasheet.

Crosstalk and phase offset: A crosstalk of only -74 dB or less is observed, even for directly adjacent channels. This is close to the reported crosstalk of the Intan chip RHA2132 (typically -78 dB). Thus, our measurements indicate that the design is robust and does not include any significant sources of crosstalk, besides the internal amplifier channel crosstalk.

Figure 8 shows the phase difference between a 1 kHz reference signal from the function generator and the signal measured with the acquisition system. A delay of 0.48 ms corresponding to a phase shift θ of 172.8° was observed. The lagging time from the input to the output is similar to the reported phase responses of the amplifier RHA2132.

Settling time: We have observed a duration of $\sim$450 μs for the channels to return back to the baseline. This is in agreement with the recommended duration of a fast settle pulse, i.e., $2.5/f_H$ (Intan technologies html) or 500 μs with our system (f_H = high cut-off frequency = 5 kHz). When large transient unwanted signals are encountered, an active-high digital pulse outputted on the *settle* pin resets the amplifiers to the baseline, and recordings are performed again as normal. Yet this blanking and resetting period would cause a data loss of tens of milliseconds, which may become critical in multi-channel stimulations. A mixed-signal topology as proposed in Nguyen *et al* (2013) could then be considered to recover the neural information during these stimulation pulses.

4.2. Multiplexed, large-scale data acquisition

Having demonstrated the low-noise recording characteristics of our system, we next have evaluated its ability to record neuronal action potentials. Figure 9 displays a representative segment of spontaneous firing activity recorded in the hippocampus and its surroundings in a freely moving rat. The majority of functional electrodes report action potentials, with more spiking activity on channels 25–28 which are located in the hippocampus. Tetrode 2 and 8 are non-functional because they are open (no connected wires). Their recorded waveforms, however, look different due to the couplings of nearby tetrodes and the surroundings. The microwires from the same tetrode display a similar pattern, indicating the correct signal sweeping and conversion (see figure 9).

The low amplifier noise enables reliable spike detection with peak amplitudes as low as 35 μV_{pp}. Examples of detected waveforms have been displayed in figure 3.

4.3. Signal analysis of recorded data

A section of the recorded data is shown in figure 10(A). Figure 10(B) displays three out of 19 analyzed templates. Figure 10(C) shows the data set with a finer temporal resolution.

The only pre-processing used is the band-pass filtering of the signal from 0.6 to 6 kHz (figure 11). The chosen type of filter decides the amplitude threshold in the next step since the filter type influences the gain ripple, resulting in different noise levels. While the average noise level of a 15.7-ms data sequence is $\sim$8.42 μV_{rms} with a 2 nd-order Butterworth filter, this increases to $\sim$11.16 μV_{rms} with a 2 nd-order elliptic filter.

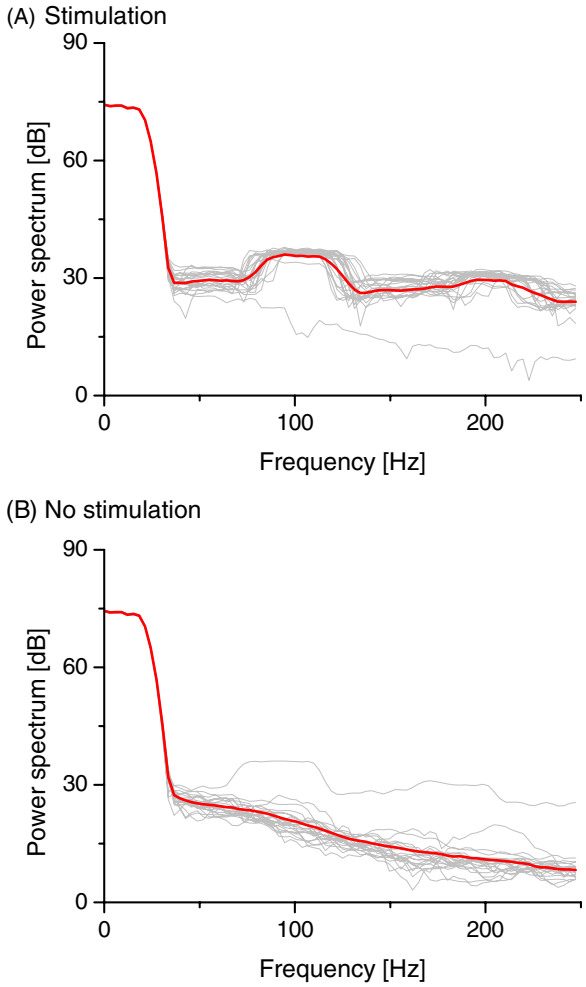

Figure 15. Power spectrum of a matrix of 216-ms data with (A) LED stimulation, and (B) no stimulation.

The locations of the Neuralynx detected spikes (×) are shown in figure 11. The spikes detected using an estimation of the adaptive median noise ($4.5\sigma_n$ or 50 ± 1.2 μV in this case) are indicated by the '+' symbol (figure 11). The differences between the two detection methods are clearly visible with overlapping spikes or small-amplitude signals. The performance of the detection procedure for different threshold levels is summarized in table 2. The number of false positives is relatively high ($\geqslant$37.5%) compared to the spikes detected with the Neuralynx system, indicating that more preprocessing needs to be done for a cleaner recorded data. Lowering the threshold value indeed reduces the number of missed spikes, i.e. both the detected hit rate and the sorted hit rate increase, but at the same time the number of false positives increases even more. The optimal trade-off between the number of missed spikes and the number of false positives requires further testing and optimization.

From the time stamps of each cluster, a neural pattern (template) is constructed by averaging the signals. The detected spikes are classified based on the highest correlation score with the set of predefined templates.

Table 3 shows the performance of the template matching algorithm for several clusters. The accuracy of the sorting process depends strongly on the features of the templates. We achieve more than 84% similarity between the developed processing sequences and the offline analyzed cluster when the templates are visibly different (figure 12(A)). Yet the similarity drops significantly when the templates look similar (figure 12(B)). The matching rates are in between when the templates only indicate several similarities (figure 12(C)). Additional work (offline) to achieve suitable feature extraction (e.g. principal component analysis, wavelet analysis, etc.) could be done to improve the matching accuracy. This is, however, beyond the scope of this work.

Since satisfactory results are obtained in the case of distinct templates, we decided to implement this process as a starting point toward a 'real-time' closed-loop stimulation and recording system. Thus, estimating the total processing time was the first requirement.

Figure 13(A) displays the average processing time required for each data sequence for four channels. Signal filtering is clearly the most time-consuming process, which on average amounts up to ~50% of the total processing time (for example 0.16 ms out of 0.28 ms for a 2-ms data). This value also means that we literally cannot achieve a 'real-time' performance for 32 channels within 2 ms with the existing algorithm (figure 13). A spike sorting algorithm without the filtering step would be advantageous and indispensable in that context.

We have chosen to control our application with a data length of 8 ms because this allows us to develop a more complex and accurate algorithm, as well because it offers scalability with only small modifications.

4.4. Online acquisition and spike sorting

The developed signal processing algorithm has been validated in real-time with 8 tetrodes (32 channels) every 8 ms.

Figure 14(A) illustrates the LED stimulation protocol based on the spike sorting. A TTL pulse of 200 ms is delivered to the optical fiber when cell 12 fires. Otherwise, no stimulation is performed. In the worst-case situation (the detected spike is at the first position of the acquisition), the LED is turned on at the 16th ms. The zoomed-in plot of 40-ms data within the moment of stimulation on cell 12 also reveals that the neurons do not get excited immediately when the LED is turned on, but only after a few milliseconds (figure 14(B)).

Tetrodes pick up the neural responses to the light pulse stimulation differently depending on their relative positions from the optical fiber in the brain (see figure 14(C)–(D)). The stimulation effects are much stronger on tetrodes close to the optical fiber and located in the hippocampus (figure 14(C)). With tetrodes placed far away from the optical fiber, the effects are much weaker, though still visible (see figure 14(D)).

The overall effect of stimulation on tetrode 1 (in the hippocampus) is plotted in figure 15. A stimulation matrix is constructed by taking 8 ms data of the clustered spikes (template 12) and the 208-ms data acquired after the stimulation epoch. The no-stimulation effect is built with 216-ms data of random positions. The signals are analyzed by their power spectrum in the Chronux package (Mitra and Bokil 2008). In the presence of LED stimulation, we have observed

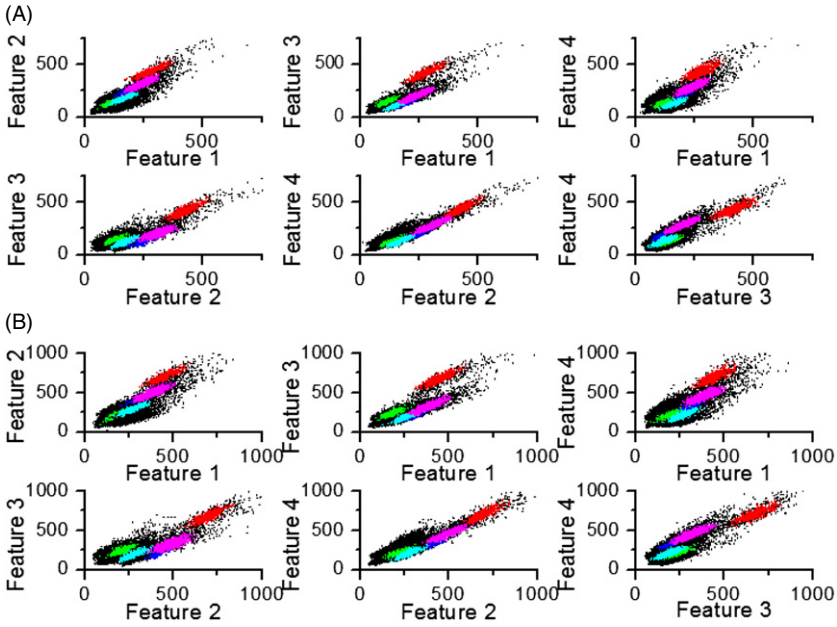

Figure 16. Examples of clustering of spikes detected in data sampled at (A) 12.5 kHz, and (B) 32 kHz. The energy landscape as well as the clustering is very similar regardless of how the spikes are sampled. Spikes automatically clustered into one of the five highest rated clusters are highlighted.

peaks in the gamma oscillation, i.e., a bump around 100 Hz and a smaller one around 200 Hz, which confirms the functionality of the implemented system.

5. Discussion

We have designed and implemented a closed-loop system capable of optical stimulation and multi-channel electrical recordings, and have demonstrated it in opto-genetically transfected rats.

The system is interfaced digitally and easy to scale up. The low sensitivity to noise and interferences has made digital transmission the choice for many designers (Mavoori *et al* 2005, Venkatraman *et al* 2009). Moreover, the Intan amplifier with integrated multiplexers reduces the number of wires by a factor of 32, thus reducing the weight of the headstage. Most control signals can be shared between multiple PCBs, i.e., the number of wires increases by just one with every additional headstage. In addition, the digital inputs can be acquired in parallel with a rate of 10 MHz. Thus, few modifications are required to expand the prototype, for recording with a larger number of channels (e.g., 96 channels).

The sampling rate of 12.5 kHz is lower than most systems currently in use, but is nonetheless enough for accurate spike sorting with standard methods (Schmitzer-Torbert *et al* 2005). Because each neuron is recorded on four channels, the most important measure for spike sorting are the amplitudes of each spike on each channel. We tested the accuracy of spike sorting 12.5 kHz recorded spikes by subsampling data recorded at 32 kHz (by Neuralynx), and then comparing the automated (offline) clustering of the subsampled and 32 kHz sampled spikes to manually clustered spikes from the same data. The automated clustering (KlustaKwik) turned out almost the same for the differently sampled data. The correct classification rate for usable manually selected clusters was 74.5% for spikes detected and sampled at 32 kHz, and 75.5% for spikes detected and sampled at 12.5 kHz. The false alarm rates for these clusters were 36.3% (32 kHz) and 35.4% (12.5 kHz). Errors are due to different filtering used by Neuralynx and our spike detection algorithm, elimination of noise in the manually clustered data, and the manual corrections that were made. Figure 16 shows how similar the results are between differently sampled data. The KlustaKwik clustering algorithm was run using two features from each channel's recording of the spikes: the energy (mean square of the waveform) of the spike, and the energy normalized first principle component of the spike waveform (Schmitzer-Torbert *et al* 2005). The energy landscape looks very similar regardless of how the spikes are sampled (compare figure 16(A) and (B)). The post-normalization first principal component (PC1) contains mostly information about the width of the spikes. There is more difference between differently sampled spikes (32 kHz sampled spikes are measured more accurately by PC1), but this information does not contribute as much to the clustering of the spikes.

Our amplifier board is able to acquire stable signals with a high signal-to-noise ratio from moving rats. However, one should try to reduce the length and weight of the connectors and cables. Since a small and light wireless recording system with the same signal quality only covers a short transmission range (∼5 m) (Fan *et al* 2011) and behavioral experiments are performed with a further distance, improvements toward minimizing the cable size and weight will be carried out in the next generation.

The implemented signal processing method has been proven to be robust and relatively reliable. An 8-ms acquisition and analysis time has been achieved for at least 32 channels on a standard workstation. Spike sorting based on template

matching has been shown to provide a basic single-unit discrimination method for non-overlapping spike events. Additional work is needed to classify newly appearing or disappearing spikes, and to analyze simultaneously firing neurons observed on the same set of channels (Segev *et al* 2004).

More complicated algorithms (e.g., principal component analysis, wavelet decomposition) may enhance the recognition of cell assemblies, and consequently improve the accuracy of spike sorting. Although the processing time of the 'offline' part of the algorithm will probably increase, the 'online' part, still being made of simple operations, most likely will not inflate time. Modern techniques, such as a GPU (graphical processing unit) (Hafeez *et al* 2012) and FPGA (Chen *et al* 2008), could also be used to improve the processing speed.

We have implemented the controlling software in LabVIEW, a user-friendly interface. The acquisition and signal analysis are multi-threaded, ensuring the real-time requirements. The compatibility of the card driver and LabVIEW (both from National Instruments) make the implementation easier.

However, the integration between Matlab and Labview is not straightforward. The Mathscript node, a platform written by Labview for Matlab code, is significantly slower than the Matlab platform itself. To keep up the high speed, the Matlab code has to be converted into a C DLL to be imported in Labview. A wrapper code has specifically been written for this purpose.

A potential improvement would be to write the spike sorting algorithm in C or Java language, rather than an interpreted language, such as the Matlab used in this project. These compiled programming languages are believed to retain or improve not only the speed compatibility but also the multi-threading capabilities for the analysis. Yet, the final decision depends on the trade-offs between programming efficiency and execution efficiency.

6. Conclusion

We have developed an optical stimulation—electrical recording system, integrating template matching for online spike sorting, to demonstrate closed-loop operation. Although various neural stimulation systems have been developed before, only very few allow closed-loop operation based on online spike sorting. The present device offers further design possibilities and flexibility for controlling the stimulation selectivity by the implementation of the template matching algorithm. The processing of data sequences of 8 ms, including acquisition, detection, classification and stimulation, has successfully been validated in awake rats, indicating a significant improvement compared to existing 'real-time' systems. Adding more features to improve the spike sorting accuracy is an important work for the future.

Acknowledgments

This work has been done in the frame of the FP7 ENLIGHTENMENT project (284801). The project ENLIGHTENMENT acknowledges the financial support of the Future and Emerging Technologies (FET) programme within the Seventh Framework Programme for Research of the European Commission.

References

Q3

Cerf M, Thiruvengadam N, Mormann F, Kraskov A, Quiroga R Q, Koch C and Fried I 2010 On-line, voluntary control of human temporal lobe neurons *Nature* **467** 1104–10

Chandra R and Optican L 1997 Detection, classification, and superposition resolution of action potentials in multiunit single-channel recordings by an on-line real-time neural network *IEEE Trans. Biomed. Eng.* **44** 403–12

Chen Y-C, Lee Y-T, Yeh S-R and Chen H 2008 A bidirectional, flexible neuro-electronic interface employing localised stimulation to reduce artifacts *Proc. 4th Int. IEEE/EMBS Conf. on Neural Engineering* pp 46–50

Cunningham J P, Nuyujukian P, Gilja V, Chestek C A, Ryu S I and V Shenoy K 2010 A closed-loop human simulator for investigating the role of feedback-control in brain-machine interfaces *J. Neurophysiol.* **105** 1932–49

Fan D *et al* 2011 A wireless multi-channel recording system for freely behaving mice and rats *PloS ONE* **6** e22033

Fountas K N and Smith J R 2007 A novel closed-loop stimulation system in the control of focal, medically refractory epilepsy *Acta Neurochirurgica Suppl.* **97** 357–62

Franke F, Natora M, Boucsein C, J. Munk M H and Obermayer K 2010 An online spike detection and spike classification algorithm capable of instantaneous resolution of overlapping spikes *J. Comput. Neurosci.* **29** 127–48

Gradinaru V, Mogri M, Thompson K R, Henderson J M and Deisseroth K 2009 Optical deconstruction of Parkinsonian neural circuitry *Science* **324** 354–9

Guger C, Gener T, A. Pennartz C M, Brotons-Mas J R, Edlinger G, Badia S B, Verschure P, Schaffelhofer S and Sanchez-Vives M V 2011 Real-time position reconstruction with hippocampal place cells *Front. Neuroprosthetics* **5** 85

Hafeez A, Asghar W, Rafique M M, Iqbal S M and Butt A R 2012 Gpu-based real-time detection and analysis of biological targets using solid-state nanopores *Med. Biol. Eng. Comput.* **50** 605–15

Harris K D, Henze D A, Csicsvari J, Hirase H and Buzsaki G 2000 Accuracy of tetrode spike separation as determined by simultaneous intracellular and extracellular measurements *J. Neurophysiol.* **84** 401–14

Hashimoto T, Elder C M, Okun M S, Patrick S K and Vitek J L 2003 Stimulation of the subthalamic nucleus changes the firing pattern of pallidal neurons *J. Neurosci.* **23** 1916–23

Hauptmann C, Popovych O and Tass P A 2007 Desynchronizing the abnormally synchronized neural activity in the subthalamic nucleus: a modeling study *Expert Rev. Med. Devices* **4** 633–50

Huber D, Petreanu L, Ghitani N, Ranade S, Hromadka T, Mainen Z and Svoboda K 2008 Sparse optical microstimulation in barrel cortex drives learned behaviour in freely moving mice *Nature* **451** 61–6

Intan technologies, L. (html). Rha2000 series multi-channel bioamplifier chips

Q4

Jackson A J M and Fetz1 E E 2006 Long-term motor cortex plasticity induced by an electronic neural implant *Nature* **444** 56–60

Jensen A L and Durand D M 2009 High frequency stimulation can block axonal conduction *Exp. Neurology* **220** 57–70

Kadir S N, Goodman D F M and Harris K D 2013 High-dimensional cluster analysis with the masked em algorithm *Quantit. Biol.* (Submitted)

Kloosterman F, Davidson T J, Gomperts S N, Layton S P, Hale G, Nguyen D P and Wilson M A 2009 Micro-drive array for

chronic in vivo recording: drive fabrication *J. Visualized Experiments* e1094
Liang S-F, Liao Y-C, Shaw F-Z, Chang D-W, Young C-P and Chiueh H 2011 Closed-loop seizure control on epileptic rat models *J. Neural Eng.* **8** 045001
Mavoori J, Jackson A, Diorio C and Fetz E 2005 An autonomous implantable computer for neural recording and stimulation in unrestrained primates *J. Neurosci. Methods* **148** 71–7
Mitra P and Bokil H 2008 *Observed Brain Dynamics* (Oxford: Oxford University Press)
Nguyen T, Musa S, Eberle W, Bartic C and Gielen G 2013 Mixed signal template-based reduction scheme for stimulus artifact removal in electrical stimulation *Med. Biol. Eng.* **51** 449–58
Oliynyk A, Bonifazzi C, Montani F and Fadiga L 2012 Automatic online spike sorting with singular value decomposition and fuzzy c-mean clustering (software) *BMC Neurosci.* **13** 96
Peckham P H and Knutson J S 2005 Functional electrical stimulations for neuromuscular applications *Annu. Rev. Biomed. Eng.* **7** 4.1–4.34
Popovych O V, Hauptmann C and Tass P A 2005 Effective desynchronization by nonlinear delayed feedback *Phys. Rev. Lett.* **94** 164102
Potter S M, Wagenaar D A and DeMarse T B 2006 *Closing the loop: Stimulation Feedback Systems for Embodied mea Cultures* (Berlin: Springer)
Quiroga R Q, Nadasdy Z and Ben Shaul Y 2004 Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering *Neural Comput.* **16** 1661–87
Rosin B, Slovik M, Mitelman R, Rivlin-Etzion M, Haber S N, Israel Z, Vaadia E and Bergman H 2011 Closed-loop deep brain stimulation is superior in ameliorating parkinsonism *Neuron* **72** 370–84
Royer S, Zemelman B V, Barbic M, Losonczy A, Buzsaki G and Magee J C 2010 Multi-array silicon probes with integrated optical bers: light-assisted perturbation and recording of local neural circuits in the behaving animal *Eur. J. Neurosci.* **31** 2279–91
Rutishauser U, Schuman E M and Mamelak A N 2006 Online detection and sorting of extracellularly recorded action potentials in human medial temporal lobe recordings, in vivo *J. Neurosci. Methods* **154** 204–24
Sahin M, Durand D and Haxhiu M 2000 Closed-loop stimulation of hypoglossal nerve in a dog model of upper airway obstruction *IEEE Trans. Biomed. Eng.* **47** 919–25
Santos F J, Costa R M and Tecuapetla F 2011 Stimulation on demand: closing the loop on deep brain stimulation *Neuron* **72** 197–8
Scanziani M and Husser M 2009 Electrophysiology in the age of light *Nature* **461** 930–9 Review article
Schmitzer-Torbert N, Jackson J, Henze D, Harris K and Redish A 2005 Quantitative measures of cluster quality for use in extracellular recordings *Neuroscience* **131** 1–11
Segev R, Goodhouse J, Puchalla J and Berry II M J 2004 Recording spikes from a large fraction of the ganglion cells in a retinal patch (technical report) *Nature Neurosci.* **7** 1155–62 **Q5**
Stark E, Koos T and Buzski G 2012 Diode probes for spatiotemporal optical control of multiple neurons in freely moving animals *J. Neurophysiol.* **108** 349–63
Sunderam S, Gluckman B, Reato D and Bikson M 2010 Review: toward rational design of electrical stimulation strategies for epilepsy control *Epilepsy Behav.* **17** 6–22
Tass P A 2002 Effective desynchronization with bipolar double-pulse stimulation *Phys. Rev.* E **66** 036226
Tass P A 2003 A model of desynchronizing deep brain stimulation with a demand-controlled coordinated reset of neural subpopulations *Biol. Cybern.* **89** 81–88
Tass P A, Klosterkotter J, Schneider F, Lenartz D, Koulousakis A and Sturm V 2003 Obsessive-compulsive disorder: development of demand-controlled deep brain stimulation with methods from stochastic phase resetting *Neuropsychopharmacol* **28** (Suppl 1) 27–34
Venkatraman S, Elkabany K, Long J D, Yao Y and Carmena J M 2009 A system for neural recordings and closed-loop intracortical microstimulation in awake rodents *IEEE Trans. Biomed. Eng.* **56** 15–22
Wagenaar D A, Madhavan R, Pine J and Potter S M 2005 Controlling bursting in cortical cultures with closed-loop multi-electrode stimulation *J. Neurosci.* **25** 680–8
Wang L, Hofer B, Guggenheim J A and Povaay B 2012 Graphics processing unit-based dispersion encoded full-range frequency-domain optical coherence tomography *J. Biomed. Opt.* **17** 077007
Wilson M A and McNaughton B L 1993 Dynamics of the hippocampal ensemble code for space *Science* **261** 1055–8
Zhang J, Laiwalla F, Kim J A, Urabe H, Wagenen R V, Song Y-K, Connors B W, Zhang F, Deisseroth K and Nurmikko A V 2009 Integrated device for optical stimulation and spatiotemporal electrical recording of neural activity in light-sensitized brain tissue *J. Neural Eng.* **6** 055007
Zhang Y-P and Oertner T G 2007 Optical induction of synaptic plasticity using a light-sensitive channel *Nature Methods* **4** 139–41
Zrenner C, Eytan D, Wallach A, Thier P and Marom S 2010 A generic framework for real-time multi-channel neuronal signal analysis, telemetry control, and sub-millisecond latency feedback generation *Front. Neurosci.* **4** 173

QUERIES

Page 1

Q1

Author: Please provide complete affiliations (lab/department/center/faculty/institute/university/city/state/country).

Q2

Author: Please be aware that the color figures in this article will only appear in color in the Web version. If you require color in the printed journal and have not previously arranged it, please contact the Production Editor now.

Page 13

Q3

Author: Please check the details for any journal references that do not have a blue link as they may contain some incorrect information. Pale purple links are used for references to arXiv e-prints.

Q4

Author: Please provide the year and the URL address in reference 'Intan Technologies, html.'

Page 14

Q5

Author: Please check whether the list of authors is okay as set in reference 'Segev *et al* 2004.'